Clarifying the abstracts of systematic literature reviews*

By James Hartley, Ph.D.† Research Professor

Department of Psychology Keele University Staffordshire United Kingdom

Background: There is a small body of research on improving the clarity of abstracts in general that is relevant to improving the clarity of abstracts of systematic reviews.

Objectives: To summarize this earlier research and indicate its implications for writing the abstracts of systematic reviews.

Method: Literature review with commentary on three main features affecting the clarity of abstracts: their language, structure, and typographical presentation.

Conclusions: The abstracts of systematic reviews should be easier to read than the abstracts of medical research articles, as they are targeted at a wider audience. The aims, methods, results, and conclusions of systematic reviews need to be presented in a consistent way to help search and retrieval. The typographic detailing of the abstracts (typesizes, spacing, and weights) should be planned to help, rather than confuse, the reader.

Several books and review papers have been published over the last twenty-five years about improving the clarity of the abstracts of articles in scientific journals, including several recent studies [1–5]. Three main areas of importance have been discussed:

- the language, or the readability, of an abstract;
- the sequence of information, or the structure, of an abstract; and
- the typography, or the presentation, of an abstract. This paper considers the implications of the findings from research in each of these overlapping areas to the more specific area of writing abstracts for what are called "systematic reviews." Such reviews in medical journals typically use standard procedures for assessing the evidence obtained from separate studies for and against the effectiveness of a particular treatment. The term "systematic" implies that the authors have used a standard approach to minimizing biases and random errors and that the methods chosen for the approach will be documented in the materials and methods sections of the review. Examples of such reviews may be found in Chalmers's and Altman's text [6] and in papers published in medical journals, par-

ticularly *Evidence-Based Medicine*. Figure 1 provides a fictitious example of an abstract for such a paper.

THE LANGUAGE OF THE TEXT

Research on the readability of conventional journal abstracts suggests that they are not easy to read. Studies in this area typically use the Flesch Reading Ease (R.E.) scores as their measure of text difficulty [7]. This measure, developed in the 1940s, is based upon the somewhat over simple idea that the difficulty of text is a function of the length of the sentences in the text and the length of the words within these sentences. The original Flesch formula is that R.E. = 206.835 – 0.846w - 1.015s (where w = the average number of syllables in 100 words and s = the average number of words per sentence). The scores normally range from 0 to 100, and the lower the score the more difficult the text is to read; Table 1 gives typical examples. Today, Flesch R.E. scores accompany most computerized spell checkers, and this removes the difficulties of hand calculation; although different programs give slightly different results [8, 9].

Table 2 summarizes the Flesch scores obtained for numerous journal abstracts in seven studies. The low scores shown here support the notion that journal abstracts are difficult to read. With medical journals, in particular, this difficulty may stem partly from complex medical terminology. Readability scores such as these are widely

^{*} Based on invited presentation at the VIIth Cochrane Colloquium, Rome, Italy, October 1999.

[†] Author's address for correspondence: James Hartley, B.A., Ph.D., Department of Psychology, Keele University, Staffordshire, ST5 5BG, United Kingdom; email, j.hartley@psy.keele.ac.uk.

Figure 1

"Before" and "after" examples designed to show how differences in typography and wording can enhance the clarity of an abstract

ALMONDS AND RAISINS IN THE TREATMENT OF INFLUENZA IN ADULTS

Peach A, Apricot D, Plum P

Background: Almonds and raisins both have antiviral properties, but they are not widely used due to incomplete knowledge of their properties and concerns about possible adverse effects.

Objective: To assess the effects of almonds and raisins in the treatment of influenza in adults.

Search strategy: We searched the Cochrane Acute Respiratory Infections Group trials register (searched 15 December 1998), the Cochrane Controlled Trials Register (Cochrane Library Issue 1, 1999), Medline (January 1966 to December 1997) and reference lists of articles. We also contacted manufacturers and researchers in the field.

Selection criteria: Randomized and quasi-randomized studies comparing almonds and/or raisins with placebo, or comparing doses or schedules of almonds and/or raisins in adults with influenza.

Data collection: Two reviewers independently assessed trial quality and extracted data. Study authors were contacted for additional information.

Main results: Seventeen trials involving 689 people were included. Five trials involving 234 people compared almonds with placebo. Compared to placebo, almonds significantly shortened duration of fever by 23% (by 1.00 days, 95% confidence interval 0.73 to 1.29). Six trials involving 256 people compared raisins with placebo. Raisins also significantly shortened duration of fever by 33% compared to placebo (by 1.27 days, 95% confidence interval 0.77 to 1.77). The little data available directly comparing almonds and raisins (two trials involving 53 people) indicated that the efficacy of the two drugs intervals were very wide. Based on four trials of 73 people, central nervous system effects were significantly more common with almonds than raisins (relative risk 2.58, 95% confidence interval 1.54 to 4.33).

Reviewers' conclusions: Almonds and raisins appear to be equally effective in the treatment of influenza. Both drugs appear to be relatively well tolerated, although raisins may be safer.

ALMONDS AND RAISINS IN THE TREATMENT OF INFLUENZA (FLU) IN ADULTS

Peach A, Apricot D, Plum P

Background. Almonds and raisins both have antiviral properties. However, they are not widely used with adults because of incomplete knowledge of their properties. There are also concerns about their possible adverse effects.

Question. How effective are almonds and raisins in the treatment of flu in adults?

Search strategy. We searched the Cochrane Acute Respiratory Infections Group trials register (searched 15 December 1998), the Cochrane Controlled Trials Register (Cochrane Library Issue 1, 1999), Medline (January 1966 to December 1998), Embase (January 1985 to December 1997) and the reference lists of relevant articles. We also contacted manufacturers and researchers in the field.

Selection criteria. We selected those studies that (i) used randomized and quasi-randomized methods of comparing almonds and/or raisins with placebos, or (ii) compared doses or schedules of almonds and/or raisins in adults with flu

Data collection. Two reviewers independently assessed the quality of the trials and extracted the data. We also contacted the authors of the relevant studies for further information.

Main results. We decided to include seventeen trials involving 689 people.

Six trials, involving 256 people, compared raisins with placebos. Here raisins significantly shortened the duration of fever by 33% compared to placebos (i.e. by one and a quarter days).

Five trials, involving 234 people, compared almonds with placebos. Here almonds significantly shortened the duration of fever by 25% (i.e. by one day.)

Only two trials, involving 53 people, directly compared almonds and raisins. These trials indicated that the efficacy of the two drugs was comparable, although the range of results was very wide.

In four trials, with 73 people, central nervous system effects were significantly more common with almonds than with raisins

Reviewers' conclusions. Almonds and raisins appear to be equally effective in the treatment of flu in adults. Both drugs appear to be relatively well tolerated, although raisins may be safer.

Abstract courtesy of Philippa Middleton.

quoted, even though there is considerable debate about their validity, largely because they ignore the readers' prior knowledge and motivation [10, 11].

A second cause of difficulty in understanding text is that, although the wording may be simple and the sentences short, the concepts being described may not be understood by the reader. Thus, for example, al-

Table 1The interpretation of Flesch scores

Flesch Reading Ease score	Reading age	Difficulty level	Example for U.S. readers
90–100 80–89 70–79 60–69 50–59 30–49 0–29	10–11 years 11–12 years 12–13 years 14–15 years 16–17 years 18–20 years Graduate	Easy Fairly easy Average Fairly difficult	Reader's Digest Time Magazine U.S. News and World Report New York Times Henry James novel Corporate annual report Legal contract

Adapted with permission from Roberts et al. [56], copyright 1994, American Medical Association.

though the sentence "God is grace" is extremely readable (in terms of the Flesch), it is not easy to explain what it actually means! In systematic reviews, to be more specific, the statistical concepts of the confidence interval and the adjusted odds ratio (Figure 1) may be well understood by medical researchers, but they will not be understood by all readers.

A third cause of difficulty in prose lies in the scientific nature of the text that emphasizes the use of the third person, together with the passive rather than the active tense. Graetz writes of journal abstracts:

The abstract is characterized by the use of the past tense, the third person, passive, and the non-use of negatives. . . . It is written in tightly worded sentences, which avoid repetition, meaningless expressions, superlatives, adjectives, illustrations, preliminaries, descriptive details, examples, footnotes. In short it eliminates the redundancy which the skilled reader counts on finding in written language and which usually facilitates comprehension. [12]

In systematic reviews, it is easy to find sentences like "Trial eligibility and quality were assessed" that

Table 2
Flesch Reading Ease scores reported in previous research on abstracts in journal articles

Study	Source and (number of abstracts)	Mean Flesch score	Method of calculation
Dronenberger and Kowitz [55]	ERIC database (48)	55*	Hand calculation (?)
Hartley [2]	B J Ed Psychol (14)	26	IBM Xenith
Hartley [2]	Revisions of the above (14)	34	IBM Xenith
Hartley and Benjamin [26]	B J Ed Psychol (30)	21	Grammatik 5.0
Hartley and Benjamin [26]	Revisions of the above (30)	25	Grammatik 5.0
Roberts, Fletcher, and Fletcher [56]	Ann Intern Med (100)	28	Right Writer 4.0
Tenopir and Jacso [31]	APA PsycLit (300+)	19	Grammatik 4.0

^{*} Estimate from mean grade levels reported.

Note: The higher the Flesch score, the more readable the text.

would be more readable if they were written as "We assessed the eligibility and the quality of the trials." Furthermore, there are often short telegrammatic communications, some of which contain no verbs. Figure 1 provides an example (under the subheading "Selection criteria").

There are, of course, numerous guidelines on how to write clear abstracts and more readable medical text [13–16] but, at present, there are few such guidelines for writing the abstracts of systematic reviews. Mulrow, Thacker, and Pugh [17] provide an excellent early example, and there are now regularly updated guidelines in the *Cochrane Handbook* [18].

Nonetheless, even when such guidelines are followed, evaluating the clarity of medical text is not easy. But some methods of doing so may be adapted from the more traditional literature on text evaluation. Schriver, for example, describes three different methods of text evaluation—text-based, expert-based, and reader-based methods [19]:

- Text-based methods are ones that can be used without recourse to experts or to readers. Such methods include computer-based readability formulae (such as the Flesch measure described above) and computerbased measures of style and grammar.
- Expert-based methods are ones that use experts to make assessments of the effectiveness of a piece of

Table 3Differences between general practitioners (GPs) and patients in their views about particular patient information leaflets

Question	% of GPs respond- ing no	% of patients respond- ing yes
Is there anything you feel essential to include but is omitted?	80	75
Is there anything you feel should be left out but is included?	86	46
Is there anywhere where you feel the style of the language is not appropriate (e.g., patronizing or confusing)?	86	50

Data adapted from Wilson et al. [22] with permission of the authors.

text. Medical experts may be asked, for example, to judge the suitability of the information contained in a patient information leaflet.

Reader-based methods are ones that involve actual readers in making assessments of the suitability of the text, for themselves and for others. Patients, for example, may be asked to comment on medical leaflets or be tested on how much they can recall from them.

Although all three methods of evaluation are useful, especially in combination, this writer particularly recommends *reader-based methods* for evaluating the readability of abstracts in systematic reviews. This recommendation is because the readers of such systematic reviews are likely to be quite disparate in their aims, needs, and even in the languages that they speak. As the 1999 *Cochrane Handbook* put it:

Abstracts should be made as readable as possible without compromising scientific integrity. They should primarily be targeted to health care decision makers (clinicians, consumers, and policy makers) rather than just researchers. *Terminology should be reasonably comprehensible to a general rather than a specialist medical audience* [emphasis added]. [20]

Expert-based measures on their own may be misleading. For instance, there is evidence to suggest that the concerns of professionals are different from those of other personnel [21]. Wilson et al. [22], for instance, report wide differences between the responses of general practitioners (GPs) and patients in the United Kingdom in responses to questions concerning the content and usefulness of several patient information leaflets. Table 3 shows some of their replies.

THE STRUCTURE OF THE TEXT

In recent times, particularly in the medical field, there has been great interest in the use of so-called "structured abstracts"—abstracts that typically contain subheadings, such as "background," "aims," "methods," "results," and "conclusions." Indeed, the early rise in the use of such abstracts was phenomenal [23], and it has no doubt continued to be so up to the present day.

Table 4Different numbers of subheadings used in abstracts in the same volume of the *Journal of the American Medical Association*

Context Objective(s) Design Mean outcome measures Results Conclusions	Context Objective(s) Design Setting(s) Mean outcome measures Results Conclusions	Context Objective(s) Design Setting and pa- tients Mean outcome measures Results Conclusions	Context Objective(s) Design Setting Participants/sub- jects/patients Interventions Mean outcome measures Results
(6 subheadings)	(7 subheadings)	(7 subheadings)	Conclusions (9 subheadings)

Evaluation studies have shown that structured abstracts are more effective than traditional ones, particularly in the sense that they contain more information [24–31]. However, a caveat here is that some authors still omit important information, and some still include information in the abstract that does not match exactly what is said in the article [32–35].

Additional research has shown that structured abstracts are sometimes easier to read and to search than are traditional ones [36, 37], but others have questioned this conclusion [38, 39]. Nonetheless, in general, both authors and readers apparently prefer structured to traditional abstracts [40–42]. The main features of structured abstracts that lead to these findings are that:

- the texts are opened-up and clearly subdivided into their component parts, which helps the reader perceive their structure;
- the abstracts sequence their information in a consistent order under consistent subheadings, which facilitates search and retrieval; and
- the writing under these subheadings ensures that authors do not miss out anything important.

Nonetheless, there are some difficulties—and these difficulties become more apparent after considering the structured abstracts of systematic reviews. First of all, the typographic practice of denoting the subheadings varies from journal to journal [43, 44]. Second, and of more relevance here, there is a range of subheadings used both *within* and *among* journals [45, 46], which militates against rapid retrieval. Table 4 shows an example of these variations by listing the subheadings used in the abstracts in just one volume of the *Journal of the American Medical Association*. Finally, it appears that some authors omit important subheadings or present them in a different order (e.g., reporting the conclusions before the results) [47].

The implications of these difficulties are that a decision needs to be made, based upon appropriate evaluation studies, about what are the key subheadings that can be used consistently in systematic reviews. The journal *Evidence-Based Medicine*, for example, uses

the following six subheadings: "Question(s)," "Data sources," "Study selection," "Data extraction," "Main results," and "Conclusions," but the *Cochrane Hand*book [48] recommends another seven: "Background," "Objectives," "Search strategy," "Selection criteria," "Data collection and analysis," "Main results," and "Reviewers' conclusions." Presumably, these different sets of subheadings have developed over time with experience. For example, "Objective(s)" initially preceded "Question(s)" in Evidence-Based Medicine. In the future, refining these subheadings further may be possible by using appropriate typographic cueing, to separate important from minor subheadings, such as those headings used in the Journal of the American Medical Association. It will be essential, however, to use consistent terminology throughout the literature to aid both the creation of and retrieval from the abstracts of systematic reviews. Editors may consult their readers and their authors for possible solutions to this prob-

THE TYPOGRAPHIC SETTING FOR ABSTRACTS OF SYSTEMATIC REVIEWS

Early research on the typographic setting of structured abstracts in scientific articles suggests that the subheadings should be printed in bold capital letters with a line space above each subheading [49]. But this research has been done with structured abstracts that only have four subheadings. However, the abstracts of systematic reviews are likely to have more than four-subheadings—indeed, as noted above, six or seven seem typical. Also, some of these subheadings may be more important than others.

Generally speaking, there are two ways of clarifying the structure in typography. One is to vary the *typography*, the other to vary the *spacing* [50, 51]. In terms of typography, not overdoing is best; there is no need to use two cues when one will do. Thus, it may be appropriate to use bold lettering for the main subheadings and italic lettering for the less important ones, without adding the additional cues of capital letters or underlining. Also, as the subheadings appear as the first word on a line, placing a line space above them enhances their effectiveness, so there is no need to indent the subheadings as well. The abstracts published in the Cochrane Library follow this procedure.

Finally in this section, it should be noted that it is easier to read an abstract:

- 1. that is set in the same type-size (or larger) than the body of the text of the review, unlike many journal abstracts, [52];
- 2. that does not use "fancy" typography or indeed bold or italic for its substantive text [53]; and
- 3. that is set in "unjustified text," with equal word spacing and a ragged right-hand margin, rather than

in "justified text," with unequal word spacing and straight left- and right-hand margins. This is particularly the case if the abstract is being read on screen [54].

CONCLUSIONS

The research reviewed above suggests that, in presenting the abstracts to systematic reviews, attention needs to be paid to their language, their structure, and their typographic design. Figure 1 shows a "before and after" example for a fictitious abstract for a systematic review. The purpose of this example is to encapsulate the argument of this paper and to show how changes in wording and typography can enhance the clarity of an abstract for a systematic review.

ACKNOWLEDGMENTS

The author is indebted to Iain Chalmers, Philippa Middleton, Mark Starr, and anonymous referees for assistance in the preparation of this paper.

REFERENCES

- 1. BERKENKOTTER C, HUCKIN TN. Genre knowledge in disciplinary communication. Hillsdale, NJ: Erlbaum, 1995.
- 2. Hartley J. Three ways to improve the clarity of journal abstracts. Brit J Educ Psychol 1994 Jun;64(2):331–43.
- 3. SWALES JM. Genre analysis: English in academic and research settings. Cambridge, U.K.: Cambridge University Press, 1990.
- 4. SWALES JM, FEAK C. Academic writing for graduate students. Michigan: University of Michigan Press, 1994.
- 5. Anonymous. A proposal for more informative abstracts in clinical articles. Ad Hoc Working Group for Critical Appraisal of the Medical Literature. Ann Intern Med 1987 Apr; 106(4):598–604.
- 6. Chalmers I, Altman DG, eds. Systematic reviews. London, U.K.: British Medical Journal Publishing Group, 1995.
- 7. FLESCH R. A new readability yardstick. J Appl Psychol 1948 Jun;32:221–3.
- 8. MAILLOUX SL, JOHNSON ME, FISHER DG, PETTIBONE TJ. How reliable is computerized assessment of readability? Computers and Nursing 1995 Sep-Oct;13(5):221–5.
- 9. SYDES M, HARTLEY J. A thorn in the Flesch: observations on the unreliability of computer-based readability formulae. British Journal of Educational Technology 1997 Apr;28(2): 143–5.
- 10. DAVISON A, GREEN G, EDS. Linguistic complexity and text comprehension: readability issues reconsidered. Hillsdale, NJ: Erlbaum, 1988.
- 11. HARTLEY J. Designing instructional text. 3d ed. East Brunswick, NJ: Nichols, 1994.
- 12. Graetz N. Teaching EFL students to extract structural information from abstracts. In: Ulijn JM, Pugh AK, eds. Reading for professional purposes. Leuven, Belgium: ACCO, 1985:125.
- 13. NATIONAL INFORMATION STANDARDS ORGANIZATION.

- An American national standard: guidelines for abstracts. Bethesda, MD: NISO Press, 1997.
- 14. HALL GM. How to write a paper. London, U.K.: BMJ Pub Group, 1994.
- 15. HARTLEY J. What does it say? text design, medical information and older readers. In Park DC, Morrell RW, Shifren K, eds. Processing of medical information in aging patients. Mahwah, NJ: Erlbaum, 1999:233–47.
- 16. MATTHEWS JR. Successful scientific writing: a step-bystep guide for biological and medical scientists. 2d ed. New York, NY: Cambridge University Press, 2000.
- 17. Mulrow CD, Thacker SB, Pugh JA. A proposal for more informative abstracts of review articles. Ann Intern Med 1988 Apr:108(4):613–5.
- 18. CLARKE M, OXMAN AD, EDS. Cochrane reviewers' handbook 4.0 [updated July 1999]. In: The Cochrane Library, issue 2 [database on CDROM]. The Cochrane Collaboration. Oxford, U.K.: Update Software, 2000.
- 19. SCHRIVER KA. Evaluating text quality: the continuum from text-focused to reader-focused methods. IEEE Trans Prof Comm 1989 Dec;32(4):238–55.
- 20. Clarke, op. cit.
- 21. BERRY DC, MICHAS IC, GILLIE T, FORSTER M. What do patients want to know about their medicines, and what do doctors want to tell them? a comparative study. Psychol and Health 1997;12(4):467–80.
- 22. WILSON R, KENNY T, CLARK J, MOSELEY D, NEWTON L, NEWTON D, PURVES I. Ensuring the readability and efficacy of patient information leaflets. Newcastle, U.K.: Sowerby Health Centre for Health Informatics, Newcastle University. Prodigy Publication 1998; no. 30.
- 23. HARBOURT AM, KNECHT LS, HUMPHREYS BL. Structured abstracts in MEDLINE 1989–91. Bull Med Libr Assoc 1995 Apr;83(2):190–5.
- 24. Mulrow, op. cit.
- 25. HARTLEY J. Applying ergonomics to Applied Ergonomics: using structured abstracts. Appl Ergonom 1999 Dec; 30(6):535–41.
- 26. HARTLEY J, BENJAMIN M. An evaluation of structured abstracts in journals published by the British Psychological Society. Brit J Educ Psychol 1998 Sep;68(3):443–56.
- 27. HAYNES RB. More informative abstracts: current status and evaluation. J Clin Epidemiol 1993 Jul;46(7):595–7.
- 28. McIntosh N, Duc G, Sedin G. Structure improves content and peer review of abstracts submitted to scientific meetings. European Science Editing 1999 Jun;25(2):43–7.
- 29. SCHERER RW, CRAWLEY B. Reporting of randomized clinical trial descriptors and the use of structured abstracts. JAMA 1998 July 15;280(3):269–72.
- 30. TADDIO A, PAIN T, FASSOS FF, BOON H, ILERSICH AL, EINARSON TR. Quality of nonstructured and structured abstracts of original research articles in the British Medical Journal, the Canadian Medical Association Journal and the Journal of the American Medical Association. Can Med Assoc J 1994 May 15;150(10):1611–5.
- 31. TENOPIR C, JACSO P. Quality of abstracts. Online 1993 May;17(3):44–55.
- 32. FROOM P, FROOM J. Deficiencies in structured medical abstracts. J Clin Epidemiol 1993 Jul;46(7):591–4.
- 33. PITKIN RM, BRANAGAN MA. Can the accuracy of abstracts be improved by providing specific instructions? a randomized controlled trial. JAMA 1998 Jul 15;280(3):267–9.

- 34. PITKIN RM, BRANAGAN MA, BURMEISTER L. Accuracy of data in abstracts of published research articles. JAMA 1999 Mar 24–31;281(12):1110–1.
- 35. HARTLEY J. Are structured abstracts more/less accurate than traditional ones? a study in the psychological literature. Journal of Information Science, 2000;26(4).
- 36. HARTLEY J, SYDES M. Are structured abstracts easier to read than traditional ones? J Res Reading 1997;20(2):122–36. 37. HARTLEY J, SYDES M, BLURTON A. Obtaining information accurately and quickly: are structured abstracts more efficient? J Info Sc 1996;22(5):349–56.
- 38. BOOTH A, O'ROURKE AJ. The value of structured abstracts in information retrieval from MEDLINE. Health Libr Rev 1997;14(3):157–66.
- 39. O'ROURKE AJ. Structured abstracts in information retrieval from biomedical databases: a literature survey. Health Informatics J 1997 Jan;3(1):17–20.
- 40. HARTLEY, BENJAMIN, Applying ergonomics to Applied Ergonomics, op. cit.
- 41. HARTLEY, An evaluation of structured abstracts in journals published by the British Psychological Society, op. cit. 42. HAYNES, op. cit.
- 43. Hartley J, Sydes M. Which layout do you prefer? an analysis of readers' preferences for different typographic layouts of structured abstracts. J Info Sci 1996;22(1):27–37.

- 44. Hartley J. Typographic settings for structured abstracts. J Technical Writing and Communication 2001: in press.
- 45. HARTLEY, BENJAMIN, An evaluation of structured abstracts in journals published by the British Psychological Society, op. cit.
- 46. HARTLEY, Which layout do you prefer?, op. cit.
- 47. O'ROURKE, op. cit.
- 48. Clarke, op. cit.
- 49. HARTLEY, Which layout do you prefer?, op. cit.
- 50. HARTLEY, Designing instructional text, op. cit.
- 51. HARTLEY, Typographic settings for structured abstracts, op. cit.
- 52. HARTLEY, Three ways to improve the clarity of journal abstracts, op. cit.
- 53. HARTLEY, Designing instructional text, op. cit.
- 54. IBID
- 55. Dronberger GB, Kowitz GT. Abstract readability as a factor in information systems. J Am Soc Inf Sci 1975;26:108–11.
- 56. ROBERTS JC, FLETCHER CH, FLETCHER SW. Effects of peer review and editing on the readability of articles published in Annals of Internal Medicine. JAMA 1994 Jul 13;272(2):119–21

Received February 2000; accepted May 2000